

A Modified Weibull Distribution

C. D. Lai, Min Xie, *Senior Member, IEEE*, and D. N. P. Murthy, *Member, IEEE*

Abstract—A new lifetime distribution capable of modeling a bathtub-shaped hazard-rate function is proposed. The proposed model is derived as a limiting case of the Beta Integrated Model and has both the Weibull distribution and Type I extreme value distribution as special cases. The model can be considered as another useful 3-parameter generalization of the Weibull distribution. An advantage of the model is that the model parameters can be estimated easily based on a Weibull probability paper (WPP) plot that serves as a tool for model identification. Model characterization based on the WPP plot is studied. A numerical example is provided and comparison with another Weibull extension, the exponentiated Weibull, is also discussed. The proposed model compares well with other competing models to fit data that exhibits a bathtub-shaped hazard-rate function.

Index Terms—Bathtub shape, hazard-rate function, modeling data, parameter estimation, Weibull distribution, Weibull probability plot.

ACRONYMS¹

Cdf	cumulative distribution function
hrf	hazard-rate function
MLE	maximum likelihood estimate
NMWD	new modified Weibull distribution
pdf	probability density function
s-	implies the statistical meaning
Sf	survival function
WPP	Weibull probability paper.

NOTATION

a, b	NMWD parameters, $a > 0, b > 0$
$h(t)$	hrf
$H(t)$	cumulative hrf
t	time
t^*	change point of $h(t)$
x	$\log(t)$
y	$\log[-\log(\bar{F}(t))]$
β	Weibull shape parameter, $\beta > 0$
λ	Weibull scale parameter, $\lambda > 0$.

Manuscript received November 1, 1999; revised March 22, 2000, April 30, and October 1, 2001. Responsible Editor: J. A. Nachlas.

C. D. Lai is with the Institute for Information Sciences and Technology, Massey University, Palmerston North, New Zealand (e-mail: C.Lai@massey.ac.nz).

M. Xie is with the Department of Industrial and Systems Engineering, National University of Singapore, Kent Ridge Crescent 119260, Singapore (e-mail: Xie@ieee.org).

D. N. P. Murthy is with the Department of Mechanical Engineering, The University of Queensland Q4072, Australia (e-mail: Murthy@mech.uq.edu.au).

Digital Object Identifier 10.1109/TR.2002.805788

¹The singular and plural of an acronym are always spelled the same.

I. INTRODUCTION

MANY lifetime distributions have bathtub-shaped hrf and many real-life data exhibit this property. A state-of-the-art survey on the class of such distributions is in [9]. This paper studies NMWD that generalizes the traditional Weibull distribution.

The usual 2-parameter Weibull distribution can be specified through its Sf:

$$\bar{F}(t; \beta, \lambda) = \exp[-(\lambda \cdot t)^\beta]. \quad (1)$$

This usual distribution does not have a bathtub hrf. However, several distributions derived from the basic Weibull distribution have bathtub-shaped hrf. Two of these are: i) the exponentiated Weibull distribution proposed in [10]; ii) the competing risk and sectional models involving 2 Weibull distributions, such as those discussed in [5].

This paper examines the statistical inference aspects and the modeling of a given data-set by the NMWD.

- Section II gives the details and some properties of NMWD.
- Section III discusses the relationship of NMWD to other well-known distributions.
- Section IV looks at the issue of determining when a given data-set can be modeled by NMWD. The question of applicability is resolved by using WPP plots.
- Section V estimates the parameters of NMWD.
- Section VI gives an illustrative example where the data are modeled by NMWD.

II. THE NEW MODIFIED WEIBULL MODEL

The lifetime distribution for NMWD arises from taking appropriate limits on the Beta integrated distributions from [8]. The Sf is

$$\bar{F}(t) = \exp[-a \cdot t^b \cdot \exp(\lambda \cdot t)] \quad (2)$$

with parameters $a > 0, b \geq 0$, and $\lambda > 0$.

The pdf and hrf are:

$$f(t) = a \cdot (b + \lambda \cdot t) \cdot t^{b-1} \cdot \exp(\lambda \cdot t) \cdot \exp[-a \cdot t^b \cdot \exp(\lambda \cdot t)], \quad (3)$$

$$h(t) = a \cdot (b + \lambda \cdot t) \cdot t^{b-1} \cdot \exp(\lambda \cdot t). \quad (4)$$

A. Characterization of $h(t)$

The shape of $h(t)$ depends only on b because of the factor, t^{b-1} , and because the remaining 2 parameters have no influence.

Cases i and ii are considered:

Case i: $b \geq 1$: Equation (4) shows that:

- 1) $h(t)$ is increasing in t , implying an increasing hrf.
- 2) $h(0) = 0$ if $b > 1$; $h(0) = a \cdot b$ if $b = 1$.

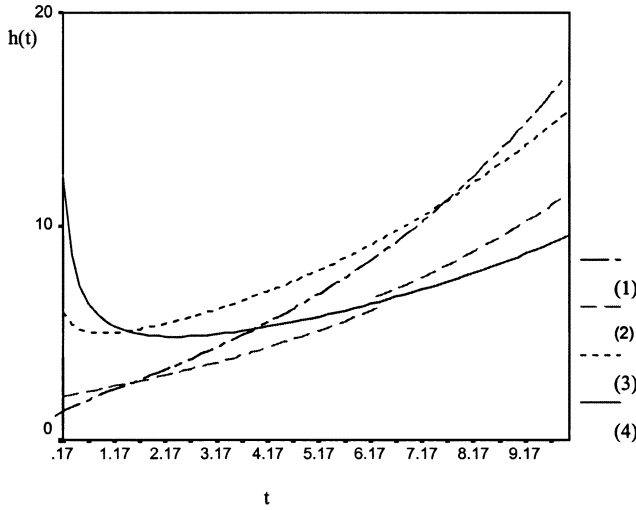


Fig. 1. hrf for the NMWD [(1) $b = 1.4, a = 1$, (2) $b = 1, a = 2$, (3) $b = 0.8, a = 5$, (4) $b = 0.4, a = 10$].

3) $h(t) \rightarrow \infty$ as $t \rightarrow \infty$.

Case ii: $0 < b < 1$: Thus:

1) $h(t)$ initially decreases and then increases in t , implying a bathtub shape for it.

2) $h(t) \rightarrow \infty$ as $t \rightarrow 0$, and $h(t) \rightarrow \infty$ as $t \rightarrow \infty$.

3) The derivative of $h(t)$ intersects the t axis only once, at t^* for $t > 0$. $h(t)$ is decreasing for $t < t^*$, and is increasing for $t > t^*$ which is given by

$$t^* = \frac{\sqrt{b} - b}{\lambda}. \quad (5)$$

The interesting feature is that t^* decreases as λ increases. This is interpreted as follows. The $h(t)$ for the Weibull distribution can be constant, increasing, or decreasing. The $h(t)$ for the new model differs from that, due to additional factor $\exp(\lambda \cdot t)$ which can be viewed as an acceleration factor increasing in both t and λ . As a result, t^* decreases as λ increases.

Fig. 1 plots $h(t)$ for 4 combinations of a and b , with $\lambda = 0.1$. The values of a, b are selected to indicate the 2 possible different shapes (increasing and bathtub) for $h(t)$.

III. RELATIONSHIP TO OTHER DISTRIBUTIONS

NMWD is related to the following well-known distributions.

A. Weibull Distribution

For $\lambda = 0$ in (2), NMWD reduces to

$$\bar{F}(t) = \exp(-a \cdot t^b), \quad (6)$$

which is a common 2-parameter Weibull distribution.

B. Extreme-Value Distribution

For $b = 0$ in (2), NMWD reduces to

$$\bar{F}(t) = \exp[-a \cdot \exp(\lambda \cdot t)]. \quad (7)$$

This is a type I extreme-value distribution and is also known as a log-gamma distribution [7]. Sometimes it is referred to as a log-Weibull. Unfortunately, the type-I extreme-value distribution is not useful as a model for life distributions, because its support spreads over the whole real line.

C. Beta-Integrated Model

The beta-integrated model was first introduced in [8]. The model's cumulative hrf and Sf are:

$$H(t) = a \cdot t^b \cdot (1 - d \cdot t)^c, \quad 0 < t < 1/d; \quad (8)$$

$$\bar{F}(t) = \exp[-H(t)]; \quad a, b, d > 0; c < 0. \quad (9)$$

Set $d = 1/n$. $c = \lambda \cdot n$, For $n \rightarrow \infty$,

$$\left(1 - \frac{t}{n}\right)^{-\lambda \cdot n} \rightarrow \exp(\lambda \cdot t),$$

and this yields,

$$H(t) = a \cdot t^b \cdot \exp(\lambda \cdot t), \quad (10)$$

which is the cumulative hrf for NMWD.

Relation to the Hjorth Model [3]: An earlier, but important bathtub-shaped failure model was proposed in [3]. The model is essentially a competing risk model involving 2 sub-populations with one having an increasing hrf and the other having a decreasing hrf. As a result, the hrf for this Hjorth model is

$$h(t) = \delta \cdot t + \frac{\theta}{1 + \beta \cdot t}, \quad 0 < \delta < \theta \cdot \beta. \quad (11)$$

On the other hand, the hrf of NMWD is given by (4).

For $\lambda = 0$, NMWD becomes a Weibull model. If $b = 2$ in addition to $\lambda = 0$, NMWD reduces to the Rayleigh distribution. Thus NMWD is not directly related to the Hjorth model (11) except when $\theta = 0$ for which case, (11) becomes the hrf for the Rayleigh distribution.

Reference [3] obtained asymptotic results for the MLE of survival probabilities, and when possible compared them with the MLE based on the Weibull, Rayleigh, and exponential distributions. In contrast, the main emphasis in this paper is on the graphical aspects of the model, such as the WPP plot.

IV. MODELING A DATA SET

This section discusses the problem of determining whether a given data set can be adequately modeled by NMWD. There are basically 2 approaches to determining if a given data-set can be adequately modeled by a given distribution:

- 1) Graphical
- 2) Statistical.

This section confines attention to the graphical approach. Because NMWD is related to the Weibull distribution, the appropriate graphical method is the WPP plot.

A. WPP Plot for NMWD

The WPP plot involves plotting the variable y versus another variable x , where these are related to the Cdf and t by the Weibull transformation:

$$y = \log[-\log(\bar{F}(t))], \quad \text{and} \quad x = \log(t).$$

Let C_1 denote this plot. If the Cdf is given by (6), then the plot is a straight line. If not, it is a curve. Reference [4] deals with the WPP plots for a mixture of 2 Weibull distributions; [5] deals with WPP plots for the competing-risk and sectional models; and [6] deals with the WPP plot for the exponentiated Weibull distribution. The C_1 for NMWD is:

$$y = \log(a) = b \cdot x + \lambda \cdot \exp(x), \quad -\infty < x < \infty. \quad (12)$$

This is a nonlinear relationship; Fig. 2 is a typical plot for it.

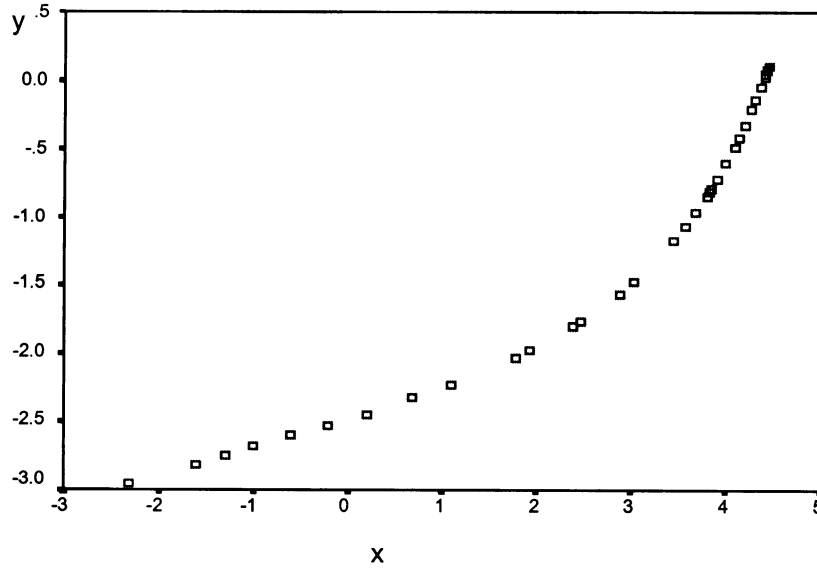


Fig. 2. A typical WPP plot for the NMWD.

Characterization of C_1 : 1) The curve C_1 is convex: the line segment that joins any 2 points on the curve does not lie below the curve.

2) The asymptotes of C_1 are:

- As $x \rightarrow -\infty$ (or $t \rightarrow 0$), an approximation to (12) is the straight line $y = \log(a) + b \cdot x$.
- As $x \rightarrow \infty$ (or $t \rightarrow \infty$), an approximation to (12) is $y \approx \exp(x) = t$.

3) Let x_0 and y_0 denote the coordinates of C_1 intersecting the x (corresponding to $y = 0$) and y (corresponding to $x = 0$) axes. Then,

$$\log(a) + b \cdot x_0 + \lambda \cdot \exp(x_0) = 0, \quad y_0 = \log(a) + \lambda. \quad (13)$$

Either $x_0 < 0$ and $y_0 > 0$ or vice versa, and $b \cdot |x_0| < |y_0|$. These are important in parameter estimation as shown in Section V.

B. An Alternative Plot

An alternative plot, C_2 (not a WPP plot) is:

$$y = \log \left[-\log \left(\overline{F}(t) \right) \right], \quad \text{and} \quad x = t \quad (14)$$

$$y = \log(a) + b \cdot \log(x) + \lambda \cdot x; \quad 0 < t < \infty. \quad (15)$$

This is again a nonlinear relationship.

Characterization of C_2 : 1) The curve is concave.

2) The asymptotes of C_2 are: • As $x \rightarrow -\infty$ (or $t \rightarrow 0$), (15) can be approximated by $y \approx b \cdot \log(t)$.

- As $x \rightarrow \infty$ (or $t \rightarrow \infty$). Equation (15) can be approximated by $y \approx \lambda \cdot x$.

The shape of C_2 (for small and large t) is different from that for C_1 .

C. Modeling Data Set

Given a data set, the first problem is to determine if NMWD is appropriate or not to model the data set. The WPP plot (based on the data) is useful for this purpose. This depends on the type of failure-data available for modeling.

This paper is limited to where the data-set available for modeling is the set of n failure times $\mathbf{t} \equiv \{t_1, t_2, \dots, t_n\}$. The

extensions to other data-set types is discussed in the *Comments* at the end of this section.

The WPP plot can be obtained by plotting

$$x_i = \log(t_{(i)}) \quad \text{versus} \quad y_i = \log \left[-\log \left(1 - \frac{i}{n+1} \right) \right],$$

for $1 \leq i \leq n$;

and $\{t_{(1)} < t_{(2)} < \dots < t_{(n)}\}$ is the set obtained by ordering of the original data. One can use other forms (e.g., $(i - 0.3)/(n + 0.4)$), instead of $i/(n + 1)$ to compute the empirical Cdf. For more on this, see [11].

If a smooth fit, based on visual assessment, to the plotted WPP plot has a shape similar to C_1 , then the data can be modeled by the new distribution. If not, the new model is not appropriate for modeling the given data set.

Comments: 1) Alternatively, one can plot $\{t_{(i)}, y_i\}$ for $1 \leq i \leq n$ and then compare the smooth fit to the data plot with C_2 to decide on the appropriateness of the model for the given data set.

2) In the graphical approach, one often uses visual evaluation as the basis to decide whether the plotted WPP has a shape similar to one of the typical shapes for the model under consideration. This is a crude approach and one needs to follow it up with more refined methods (based on rigorous statistical tests) to make the final decision whether the model under consideration is appropriate or not. There are several well-known goodness-of-fit measures, e.g., Chi-squared and Kolmogorov–Smirnov tests, which are based on the empirical distribution function. Because the distribution function for this model is very simple, both of these two tests are recommended. Because this distribution is related to an extreme value distribution, the goodness of fit tests in [7, chapter 22, section 13] would apply.

3) If the data are censored (or grouped) then the plotting of the data needs to be modified because the empirical distribution function is no longer given by $i/(n + 1)$. For further details, see [11].

V. PARAMETER ESTIMATION

As for any traditional lifetime distribution, the model parameters must be estimated based on actual data. Several methods can be used and are discussed here.

A. Based on WPP Plot

Step 1) Fit a smooth curve to the transformed data.

Step 2) Estimate λ from the slope of the straight line asymptote.

Step 3) Estimate the remaining 2 parameters from the 2 intercepts given by (13).

A possible approach is simple regression analysis using (12) and estimating the parameters by least-squares fit.

An alternate is a multiple linear regression using $x_1 = x$ and $x_2 = \exp(x)$ so that the WPP plot can be represented as the linear equation:

$$y = \log(a) + b \cdot x_1 + \lambda \cdot x_2 \quad (16)$$

to estimate the parameters, noting that x_1 and x_2 are not s -independent.

B. Method of Percentiles

It is not possible to use the method of moments to estimate the parameters of the distribution because expressions for the moments are analytically intractable. Although the expressions for the percentiles are analytically tractable, suggesting that using the estimates of the 2 quartiles and median, one can obtain the parameter estimates easily. In general, the estimates based on percentiles or moments are not very efficient when compared with of the MLE.

C. MLE Method

The likelihood function is easy to derive, and for complete data, the log likelihood is:

$$\begin{aligned} l(a, b, \lambda; \mathbf{t}) &= n \cdot \log(a) + \sum_{i=1}^n \log(b + \lambda \cdot t_i) + (b-1) \cdot \log\left(\sum_{i=1}^n t_i\right) \\ &\quad + \lambda \cdot \sum_{i=1}^n t_i + a \cdot \sum_{i=1}^n (t_i^b \cdot \exp(\lambda \cdot t_i)). \end{aligned} \quad (17)$$

The partial derivative with respect to a is:

$$\frac{\partial l}{\partial a} = \frac{n}{a} - \sum_{i=1}^n t_i^b \cdot \exp(\lambda \cdot t_i). \quad (18)$$

Equate (18) to zero:

$$\hat{a} = \frac{n}{\sum_{i=1}^n t_i^b \cdot \exp(\lambda \cdot t_i)}. \quad (19)$$

By taking the partial derivatives and equating them to zero, the log likelihood function yields the estimates for b and λ from (20) and (21)

$$\begin{aligned} \frac{\partial l}{\partial b} &= \sum_{i=1}^n \frac{1}{b + \lambda \cdot t_i} + \log\left(\sum_{i=1}^n t_i\right) \\ &\quad - a \cdot b \cdot \left(\sum_{i=1}^n t_i^{b-1} \exp(\lambda \cdot t_i)\right) = 0 \end{aligned} \quad (20)$$

TABLE I
FAILURE-TIME DATA FROM AARSET [1]

0.100	7.00	36.00	67.00	84.00
0.200	11.00	40.00	67.00	84.00
1.00	12.00	45.00	67.00	84.00
1.00	18.00	46.00	67.00	85.00
1.00	18.00	47.00	72.00	85.00
1.00	18.00	50.00	75.00	85.00
1.00	18.00	55.00	79.00	85.00
2.00	18.00	60.00	82.00	85.00
3.00	21.00	63.00	82.00	86.00
6.00	32.00	63.00	83.00	86.00

TABLE II
PARAMETER ESTIMATION AND TEST OF s -SIGNIFICANCE

	Coeff.	Std. Error	t
(Constant)	-2.43500	0.079	-30.672
X	0.38900	0.043	8.932
$\exp(X)$	0.01512	0.002	6.463

The s -significance = 0.000 for all 3 rows.

$$\begin{aligned} \frac{\partial l}{\partial \lambda} &= \sum_{i=1}^n \frac{1}{b + \lambda \cdot t_i} + \sum_{i=1}^n t_i \\ &\quad - a \cdot \lambda \cdot \left(\sum_{i=1}^n t_i^b \cdot \exp(\lambda \cdot t_i)\right) = 0. \end{aligned} \quad (21)$$

These equations need to be solved numerically to estimate $\hat{\lambda}$ and \hat{b} .

VI. AN ILLUSTRATIVE EXAMPLE

This section uses the data set (representing failure times and shown in Table I) from [1] to illustrate the modeling and estimation procedure.

As the first step for modeling and analysis, a WPP is produced, and as anticipated, a straight-line fitting is not good. Hence, the Weibull distribution is not a good model for this data set; and thus other models should be used. The shape of the Weibull plot has the same as that in Fig. 2.

The smooth fit in the plot of Weibull transformed data indicates that the new Weibull distribution is an appropriate model. The parameter estimates are obtained using a regression procedure. The estimated model is

$$y = -2.435 + 0.389x + 0.01512 \exp(x). \quad (22)$$

The $(\hat{a}, \hat{b}, \hat{\lambda}) = (0.0876, 0.389, 0.01512)$.

Table II shows the results of regression analysis for parameter estimation. These results are compared with [9] in which another interesting extension of Weibull distribution (the exponentiated Weibull) is used for the same data set.

Fig. 3 shows:

- the Weibull plots for the new model, based on (22),
- the exponentiated Weibull model, based on the parameter estimates from [10],
- the plot of the transformed data.

Table III shows the descriptive statistics for the exponentiated Weibull for the 2 distributions. The MSE for NMWD is much smaller than the MSE for the exponentiated model, indicating that NMWD fits the data better.

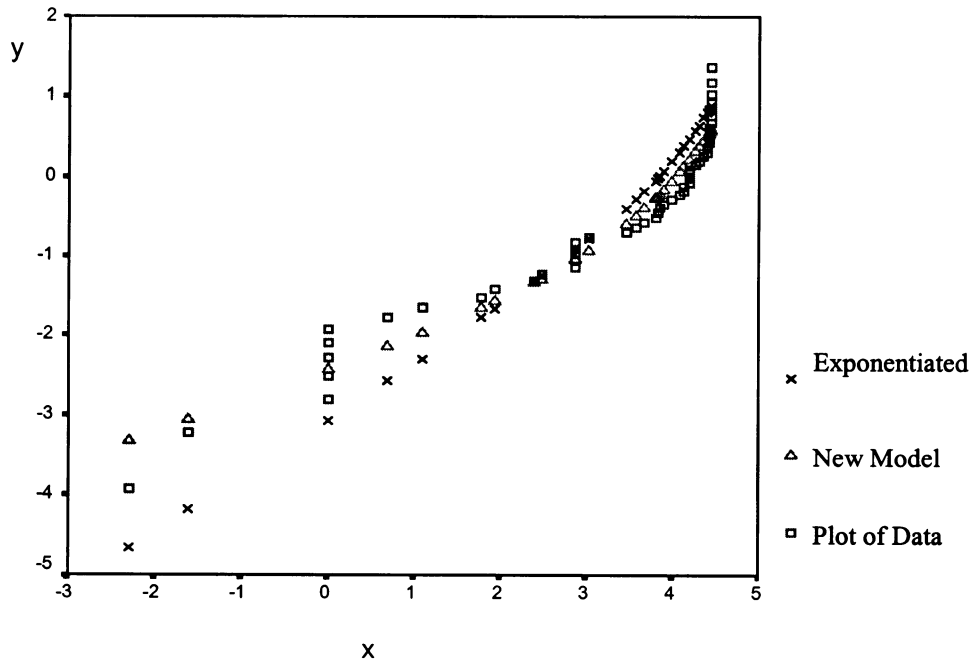


Fig. 3. Comparison of NMWD and the exponentiated Weibull for the data set from [1].

TABLE III
COMPARISON OF NMWD AND EXPONENTIATED
DISTRIBUTIONS FOR DATA FROM [1]

	<i>N</i>	Min.	Max.	Mean	Std. Dev.
ERR-NEW	50	0.00	0.60	0.07148	0.1101
ERR-MS	50	0.00	1.33	0.21900	0.2716

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor, Dr. Nachlas, for his valuable comments on an earlier version of the paper.

REFERENCES

- [1] M. V. Aarset, "How to identify bathtub hazard rate," *IEEE Trans. Rel.*, vol. R-36, no. 1, pp. 106–108, Apr. 1987.
- [2] R. C. Gupta, R. L. Gupta, and R. D. Gupta, "Modeling failure time data by Lehman alternatives," *Communications in Statistics—Theory and Methods*, vol. 27, pp. 887–904, 1997.
- [3] U. Hjorth, "A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates," *Technometrics*, vol. 22, no. 1, pp. 99–107, Feb. 1980.
- [4] R. Jiang and D. N. P. Murthy, "Graphical representation of two mixed Weibull distributions," *IEEE Trans. Rel.*, vol. R-44, no. 2, pp. 477–488, June 1995.
- [5] —, "Reliability modeling involving two Weibull distributions," *Reliab. Eng. Syst. Safety*, vol. 47, pp. 187–198, 1995.
- [6] —, "Exponentiated Weibull family: A graphical approach," *IEEE Trans. Rel.*, vol. R-48, no. 1, pp. 68–72, Mar. 1999.
- [7] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*: Wiley, 1995, vol. 2.
- [8] C. D. Lai, T. Moore, and M. Xie, "The beta integrated model," in *Proc. Int. Workshop on Reliability Modeling and Analysis—From Theory to Practice*, 1998, pp. 153–159.
- [9] C. D. Lai, M. Xie, and D. N. P. Murthy, "Bathtub shaped failure rate distributions," in *Handbook in Reliability*, N. Balakrishnan and C. R. Rao, Eds., 2001, vol. 20, pp. 69–104.

- [10] G. S. Mudholkar and D. K. Srivastava, "Exponentiated Weibull family for analysing bathtub failure rate data," *IEEE Trans. Rel.*, vol. 42, no. 2, pp. 299–302, Jun. 1993.
- [11] W. Nelson, *Lifetime Data Analysis*: Wiley, 1982.

C. D. Lai was born in Sarawak, Malaysia. He received his tertiary education in New Zealand and obtained the Ph.D. in Statistics from the Victoria University of Wellington. He held positions at the University of Auckland and the National Chiao Tung University (Taiwan) prior to coming to Massey in 1979. His research interests include statistical reliability, stochastic point processes, statistical quality control, and probability distributions. Professor Lai has published over 70 papers and coauthored 2 books.

Min Xie obtained his Licentiate in 1986 and the Ph.D. in 1987 in Quality Technology at Linköping University in Sweden. He holds the M.Sc. from the Royal Institute of Technology in Sweden. After working for 4 years at Linköping University, he joined the National University of Singapore in 1991, and was awarded the prestigious Lee Kuan Yew research fellowship. Professor Xie is an Editor of *International Journal of Reliability, Quality, and Safety Engineering* and is on the editorial board of *IIE Transactions on Quality and Reliability Engineering*. He is a Senior Member of IEEE and ASQ.

D. N. P. Murthy obtained the B.E. and M.E. from Jabalpur University and the Indian Institute of Science in India, and the M.S. and Ph.D. from Harvard University. He is the Professor of Engineering and Operations Management in the Department of Mechanical Engineering at the University of Queensland, and a Senior Scientific Advisor to the Norwegian University of Science and Technology. He has held visiting appointments at several universities in the USA, Europe, and Asia. He has authored and coauthored over 130 journal papers and 120 conference papers. He is a coauthor of 3 books: *Mathematical Modeling*, Pergamon Press, London; *Warranty Cost Analysis*, Marcel Dekker; *Reliability: Modeling, Prediction and Optimization*, Wiley; and co-editor of *Product Warranty Handbook*, Marcel Dekker. He is working on 2 books (*Weibull Models*, and *Reliability Case Studies*) to be published by Wiley. He is a member of several professional societies and is on the editorial boards of 7 international journals.

A modified Weibull distribution

Lai, C. D.

2003-03

<http://hdl.handle.net/10179/9613>

22/04/2023 - Downloaded from MASSEY RESEARCH ONLINE